

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 07-023322

(43)Date of publication of application : 24.01.1995

(51)Int.Cl.

H04N 5/765  
G11B 20/12  
G11B 31/00  
H04N 5/232  
H04N 5/91

(21)Application number : 05-147337

(71)Applicant : MATSUSHITA ELECTRIC IND CO LTD

(22)Date of filing : 18.06.1993

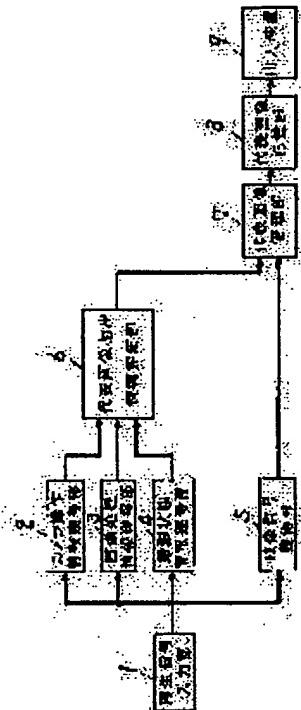
(72)Inventor : AKAHORI HIROSHI  
GOSHIMA YUKIE  
FUJIMOTO MAKOTO

## (54) MOVING IMAGE RECORDING MEDIUM, STILL IMAGE EXTRACTING DEVICE, MOVING IMAGE RECORDER, AND AUTOMATIC EXTRACTING METHOD FOR STILL IMAGE

### (57)Abstract:

PURPOSE: To automatically extract a representative image capable of sufficiently expressing the contents of a moving image by previously recording information for extracting the representative image together with a video signal in a moving image recording medium and extracting the representative image as a still image based upon the evaluation result of the representative image.

CONSTITUTION: Information recorded in the moving image recording medium is reproduced and inputted to a reproduced signal input part 1. A camera operation information acquiring part 2, a picture processing information acquiring part 3 and a photographing state information acquiring part 4 respectively read out camera operation information, picture processing information and photographing state information from the inputted reproduced signal. A representative image extracting information evaluating part 6 detects picture recording starting operation information included in the camera operation information and evaluates whether each image in the same cut is a proper image as a representative image or not until the detection of the succeeding picture recording starting operation information. A representative image managing part 7 enters a picture for one frame corresponding to the proper image evaluated as the representative image from a video signal acquiring part 5 and stores the acquired picture in a representative image storing part 8 and an output device 9 extracts and outputs the representative image.



(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平7-23322

(43)公開日 平成7年(1995)1月24日

(51)Int.Cl.<sup>6</sup> 識別記号 庁内整理番号 F I 技術表示箇所  
H 04 N 5/765  
G 11 B 20/12 103 9295-5D K  
31/00 541 L 8322-5D  
H 04 N 5/782  
7734-5C 5/91 J  
審査請求 未請求 請求項の数 7 OL (全 14 頁) 最終頁に続く

(21)出願番号 特願平5-147337  
(22)出願日 平成5年(1993)6月18日

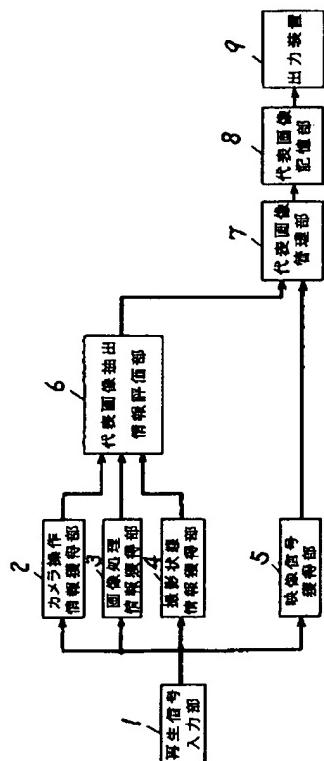
(71)出願人 000005821  
松下電器産業株式会社  
大阪府門真市大字門真1006番地  
(72)発明者 赤堀 裕志  
大阪府門真市大字門真1006番地 松下電器  
産業株式会社内  
(72)発明者 五島 雪絵  
大阪府門真市大字門真1006番地 松下電器  
産業株式会社内  
(72)発明者 藤本 風  
大阪府門真市大字門真1006番地 松下電器  
産業株式会社内  
(74)代理人 弁理士 小鶴治 明 (外2名)

(54)【発明の名称】 動画像記録媒体、静止画像抽出装置、動画像記録装置および静止画像自動抽出方法

(57)【要約】

【目的】 ビデオカメラなどで撮影された動画像の内容を十分に表現する代表画像を静止画像として自動的に抽出することとする。

【構成】 動画像を撮影する際に撮影者がカメラを操作したカメラ操作情報を取り込むカメラ操作情報獲得手段と、撮像した画像を処理して得られた画像処理情報を取り込む画像処理情報獲得手段と、センサーからの信号を処理して得られた撮影中の撮影状態情報を取り込む撮影状態情報獲得手段と、前記カメラ操作情報獲得手段からのカメラ操作情報と前記画像処理情報獲得手段からの画像処理情報と前記撮影状態情報獲得手段からの撮影状態情報の少なくとも一つをもとにして、撮影者が撮影開始操作をしてから撮影終了操作をするまでの間に撮影された動画像の中から動画像の内容を表現する代表的な静止画像を抽出する静止画像抽出手段を備える。



## 【特許請求の範囲】

【請求項1】動画像を撮影する際に撮影者がカメラを操作したカメラ操作情報を記憶するカメラ操作情報記憶部と、撮像した画像を処理して得られた画像処理情報を記憶する画像処理情報記憶部と、センサーからの信号を処理して得られた撮影中の撮影状態情報を記憶する撮影状態情報記憶部の少なくとも一つを備えることを特徴とする動画像記録媒体。

【請求項2】動画像を撮影する際に撮影者がカメラを操作したカメラ操作情報を取り込むカメラ操作情報獲得手段と、撮像した画像を処理して得られた画像処理情報を取り込む画像処理情報獲得手段と、センサーからの信号を処理して得られた撮影中の撮影状態情報を取り込む撮影状態情報獲得手段と、前記カメラ操作情報獲得手段からのカメラ操作情報と前記画像処理情報獲得手段からの画像処理情報と前記撮影状態情報獲得手段からの撮影状態情報の少なくとも一つをもとにして、撮影者が撮影開始操作をしてから撮影終了操作をするまでの間に撮影された動画像の中から少なくとも1枚の静止画像を抽出する静止画像抽出手段を備えたことを特徴とする静止画像抽出装置。

【請求項3】動画像を撮影する際に撮影者がカメラを操作したカメラ操作情報を取り込むカメラ操作情報獲得手段と、撮像した画像を処理して得られた画像処理情報を取り込む画像処理情報獲得手段と、センサーからの信号を処理して得られた撮影中の撮影状態情報を取り込む撮影状態情報獲得手段と、前記カメラ操作情報獲得手段からのカメラ操作情報と前記画像処理情報獲得手段からの画像処理情報と前記撮影状態情報獲得手段からの撮影状態情報の少なくとも一つをもとにして、撮影者が撮影開始操作をしてから撮影終了操作をするまでの間に撮影された動画像の中から少なくとも1枚の静止画像を抽出する静止画像抽出手段と、撮像装置で撮影した動画像を動画像記録媒体に記録するとともに、前記静止画像抽出手段で抽出した静止画像の情報を記録する静止画像情報記録手段を備えることを特徴とする動画像記録装置。

【請求項4】撮影者が撮影開始操作をしてから撮影終了操作をするまでの間に撮影された動画像の中から少なくとも1枚の静止画像を抽出するために、撮像装置のズームや撮影開始操作などのカメラ操作情報と撮像した画像を処理して得られた画像処理情報とセンサーからの信号を処理して得られた撮影中の撮影状態情報の少なくとも一つを入力とし、静止画像抽出知識にもとづいて各画像についての評価値を求め、評価値が高い画像を抽出することを特徴とする静止画像自動抽出方法。

【請求項5】撮影者が撮影開始操作をしてから撮影終了操作をするまでの間に撮影された動画像の中から少なくとも1枚の静止画像を抽出するために、一定条件を満足する画像に対して、撮像装置のカメラ操作情報と撮像した画像を処理して得られた画像処理情報とセンサーから

の信号を処理して得られた撮影中の撮影状態情報の少なくとも一つを入力とし、前記カメラ操作情報と前記画像処理情報と前記撮影状態情報が所定の抽出条件を満足した画像を抽出することを特徴とする静止画像自動抽出方法。

【請求項6】撮影者が撮影開始操作をしてから撮影終了操作をするまでの間に撮影された動画像の中から少なくとも1枚の静止画像を抽出するために、撮影開始操作から一定時間経過後の画像に対して、撮像装置のズームなどのカメラ操作情報と撮像した画像を処理して得られた画像処理情報とセンサーからの信号を処理して得られた撮影中の撮影状態情報の少なくとも一つを入力とし、前記カメラ操作情報と前記画像処理情報と前記撮影状態情報が所定の条件を満足した画像を抽出することを特徴とする請求項5記載の静止画像自動抽出方法。

【請求項7】撮影者が撮影開始操作をしてから撮影終了操作をするまでの間に撮影された動画像の中から少なくとも1枚の静止画像を抽出するために、撮影開始操作から撮影終了操作までの間に撮影された画像の数に一定比率を乗じた画像数以降に撮影された画像に対して、撮像装置のズームなどのカメラ操作情報と撮像した画像を処理して得られた画像処理情報とセンサーからの信号を処理して得られた撮影中の撮影状態情報の少なくとも一つを入力とし、前記カメラ操作情報と前記画像処理情報と前記撮影状態情報が所定の条件を満足した画像を抽出することを特徴とする請求項5記載の静止画像自動抽出方法。

## 【発明の詳細な説明】

## 【0001】

【産業上の利用分野】本発明は、ビデオカメラなどで撮影された動画像の中から、代表的な画像を自動的に抽出する方法及びその装置に関するものである。

## 【0002】

【従来の技術】従来の技術として、例えばVTRでは記録された動画像を管理するための情報を、ビデオテープに動画像とともに記録するというものがある。一例として、VISS (VHS Index Search System) について説明する。VISSとは、VHS方式のVTRにおいて高速頭出しを行うために開発されたものである。ビデオテープには、通常の画像情報を記録するビデオトラック以外に、この高速頭出しを行うためのVISS信号を記録するコントロルトラックが存在する。このVISS信号は、ビデオテープに画像情報を録画し始めたときに、コントロルトラックに自動的に記録される。また、ユーザが見たい場面に対してVISS信号を記録することもできる。このようにしてビデオテープ上に記録されたVISS信号を利用して、イントロサーチと呼ばれる早送り再生を行うことができる。イントロサーチとは、早送り中にVISS信号を見つけると、ある時間だけ再生状態にし、その後再び早送りす

るという動作をテープの終わりまで繰り返すものである。

### 【0003】

【発明が解決しようとする課題】しかしながら、従来のVTRにおいて自動的に記録されるViSS信号は、ビデオテープに画像情報を録画し始めたときに記録されるため、イントロサーチのようにViSS信号の付いている部分の画像を抽出した場合、抽出された画像は動画像の内容を十分に表現するものではない。また、動画像の内容を十分に表現する部分にViSS信号を付与しようとすると、ユーザが手動で行わなければならず、膨大な手間がかかる。

【0004】本発明はかかる点に鑑み、動画像の内容を十分に表現する代表画像を自動的に抽出する装置および動画像記録媒体を提供することを目的とする。

### 【0005】

【課題を解決するための手段】上記課題を解決するために、本発明の動画像記録媒体は動画像を撮影する際に撮影者がカメラを操作したカメラ操作情報を記憶するカメラ操作情報記憶部と、撮像した画像を処理して得られた画像処理情報を記憶する画像処理情報記憶部と、センサーからの信号を処理して得られた撮影中の撮影状態情報を記憶する撮影状態情報記憶部の少なくとも一つを備える。

【0006】また本発明の静止画像抽出装置は、動画像を撮影する際に撮影者がカメラを操作したカメラ操作情報を取り込むカメラ操作情報獲得手段と、撮像した画像を処理して得られた画像処理情報を取り込む画像処理情報獲得手段と、センサーからの信号を処理して得られた撮影中の撮影状態情報を取り込む撮影状態情報獲得手段と、前記カメラ操作情報獲得手段からのカメラ操作情報と前記画像処理情報獲得手段からの画像処理情報と前記撮影状態情報獲得手段からの撮影状態情報の少なくとも一つをもとに、撮影者が撮影開始操作をしてから撮影終了操作をするまでの間に撮影された動画像の中から少なくとも1枚の静止画像を抽出する静止画像抽出手段を備える。

【0007】さらに本発明の動画像記録装置は、撮像装置で撮影した動画像を動画像記録媒体に記録するとともに、前記静止画像抽出手段で抽出した静止画像の情報を記録する静止画像情報記録手段を備える。

### 【0008】

【作用】以上のような構成において、撮影時における撮影者のズームなどのカメラ操作情報や、画像処理を行って得られる例えばフォーカスやアイリスの信頼性や被写体の位置や障害物の存在状況などの画像処理情報や、センサーから得られる例えばパンなどの撮影状態情報をもとに評価を行い、評価値が高いか、もしくは一定の条件を満足する画像を動画像の中から抽出する。これによって、抽出される静止画像は動画像の内容を十分に表

現したものとなる。

【0009】また、抽出した静止画像の情報を動画像とともに記録することによって、静止画像情報をもとに代表画像を容易に検索したり、出力したりすることが可能となる。

### 【0010】

【実施例】本発明の動画像記録媒体の実施例を図1に示す。図1は動画像記録媒体の一例としてビデオテープを示すが、ビデオディスクやICメモリなどの他の記録媒体であってもよい。図1に示すように、ビデオテープには映像信号とともに、映像信号に対応してフレーム単位で代表画像の抽出情報を記録しておく。ここで代表画像の抽出情報とは、カメラ操作情報と画像処理情報と撮影状態情報である。カメラ操作とはビデオカメラで撮影した際の録画開始操作やズーム操作などであり、録画開始操作情報は録画開始操作が行われた時点にフラグを立てて録画開始点を示す情報であり、ズーム操作情報はズーム倍率を表す情報である。録画開始操作情報もズーム操作情報とともに、ビデオカメラのボタン操作をもとに検出可能な情報である。画像処理情報は像素子で撮像した映像信号をもとに自動的もしくは人間が関与して抽出処理した情報で、例えばフォーカス制御を行うために求めた映像信号の高周波成分の周波数や大きさ、あるいはフレーム間における輝度信号や色信号の差異を求めたフレーム間差分値、あるいは映像信号から被写体領域の位置や大きさなどの情報を抽出したもの、逆光や過順光の状態、さらにはγ補正值や色温度などである。撮影状態情報はカメラの撮影状態をセンサーで検出した情報で、例えば角速度センサーによって検出したパンニングなどのカメラの動き情報、あるいは光量センサーによる被写体光量や絞り開度センサーで検出したレンズの絞り開度、あるいはレンズ位置検出によるフォーカス距離などがある。

【0011】以上のような代表画像抽出情報を動画像記録媒体に備えることによって、以降の実施例で説明するように動画像中の代表画像を抽出することが可能となる。

【0012】次に本発明の静止画像抽出装置の第1の実施例について説明する。図2に本実施例の構成図を示す。本実施例は、動画像記録媒体に映像信号とともに代表画像を抽出するための情報を記録しておき、動画像記録媒体から代表画像の抽出情報を読み出して評価し、評価結果にもとづいて1カット（カメラにおいて録画開始操作をしてから録画終了操作をするまでの間に連続して撮影された動画像のかたまり）の中から代表的な画像を静止画像として抽出するものである。図2で1は再生信号入力部、2はカメラ操作情報獲得部、3は画像処理情報獲得部、4は撮影状態情報獲得部、5は映像信号獲得部、6は代表画像抽出情報評価部、7は代表画像管理部、8は代表画像記憶部、9は出力装置である。以上の

構成における各部の動作について以下で詳細に説明する。

【0013】まず、再生信号入力部1には動画像記録媒体に記録された情報を再生して入力する。カメラ操作情報獲得部2および画像処理情報獲得部3および撮影状態情報獲得部4では、再生信号入力部1に入力された再生信号からそれぞれカメラ操作情報と画像処理情報と撮影状態情報を読み出す。本実施例ではコード化された各情報をデコードする。代表画像抽出情報評価部6は、カメラ操作情報に含まれる録画開始操作情報を検出し、次の録画開始操作情報を検出するまでの同一カット内の各画像に対して、カメラ操作情報の中の録画開始操作情報以外の情報と画像処理情報と撮影状態情報をもとに、各画像がカットの代表画像として適当な画像であるかどうかの評価を行う。代表画像として適当であると評価された画像に関して、代表画像管理部7は映像信号獲得部5から1フレームの画像を取り込み、代表画像記憶部8に記憶する。出力装置9は代表画像記憶部8に記憶された代表画像を取り出して出力するものであり、ディスプレイやプリンタなどである。なお、代表画像抽出情報評価部6の動作については、後述の静止画像自動抽出方法の実施例で詳細に説明する。

【0014】以上の実施例では、動画像記録媒体に映像信号とともにあらかじめ代表画像を抽出するための情報を記録しておき、動画像記録媒体から代表画像の抽出情報を読み出して代表画像を抽出する場合について説明を行った。しかし、代表画像を抽出するための情報の一部、もしくは全部が動画像記録媒体に存在しない場合でも、動画像記録媒体に記録された映像信号を処理することによって、代表画像を抽出するための情報を獲得し、獲得した情報をもとに代表画像を抽出することができる。これについて以下の第2の実施例の静止画像抽出装置で詳細に説明する。

【0015】第2の実施例は、映像信号のみからすべての代表画像抽出情報を獲得するものである。図3に映像信号から代表画像抽出情報を獲得する装置の構成例を示す。図3で10はフレーム間差分値検出部、11はメモリ、12は変化量検出部、13はカットエンジ検出部、14はカメラワーク検出部、15は動きベクトル検出部、16はカメラワークパラメータ推定部、17は被写体情報検出部、18は動領域検出部、19は領域内特微量抽出部、20はフォーカス情報検出部、21は高域通過フィルタ、22は平均値算出部である。以上の構成における各部の動作について以下で詳細に説明する。

【0016】まず、フレーム間差分値検出部10およびカットエンジ検出部13の動作について説明する。フレーム間差分値検出部10は、動画像信号を1フレーム遅延させるためのメモリ11と、連続するフレーム間で動画像信号の差分を求める変化量検出部12からなる。動画像の連続するフレーム間の差を求める信号は輝度值

やr g b値などを用い、変化量検出部12において画素単位で連続するフレーム間の画像信号の差分演算を行い、画素ごとの差分値の総和を求めてフレーム間差分値として出力する。カットエンジ検出部13は、フレーム間差分値検出部10で求めたフレーム間差分値に対して閾値処理をする。すなわち、所定の閾値とフレーム間差分値との比較を行い、フレーム間差分値が閾値より大きい場合は、2枚のフレーム間で画像内容が大きく変化していると考えて、その部分でカットエンジがあったと判断する。ビデオカメラでは、録画開始操作を行うことによってカットエンジが生じるため、逆に画像信号からカットエンジを検出することによって録画開始操作を推定することができる。したがって、カットエンジ検出部13では、閾値を越えるフレーム間差分値が検出された時点で、録画開始操作情報を出力する。なお、図3で示したフレーム間差分値検出部10の構成は一例であり、図4で示すような他の構成でもよい。図4で44は動画像信号の1フレームにおける色ヒストグラムを求める色ヒストグラム検出部、45は求めたヒストグラムを記憶するヒストグラムメモリ、46は連続するフレーム間で色ヒストグラムの差異を検出するヒストグラム差分検出部である。図4に示す構成では、フレーム間で画素ごとの比較を行うのではなくフレーム全体で比較を行うものであるが、画面を複数のブロックに分割し、ブロック単位でフレーム間の差分を求める構成としてもよい。

【0017】次にカメラワーク検出部14について説明する。まず、動きベクトル検出部15の動作を説明する。図5は、検出する動きベクトルの画面内の位置を説明するための図である。図5(a)は、全画面で水平、垂直方向にM、N本の直線を格子状に配列した図であり、M・N個の交点は検出すべき動きベクトルの位置を示している。以下M・N個の交点を格子点と呼び、水平、垂直方向で各々 i, j 番目の格子点を、格子点 (i, j) ( $1 \leq i \leq M, 1 \leq j \leq N$ ) と呼ぶ。

【0018】本実施例で格子点位置での動きベクトルは、各格子点の周辺で複数の代表点を選び、代表点マッピングにより求める。図5(b)は、図5(a)の格子点 (i, j) 近傍を拡大した図であり、格子点とその周辺の $(2 \cdot m+1) \cdot (2 \cdot n+1)$ 個の代表点の位置関係を示す。以下、格子点 (i, j) の代表点のうち、水平、垂直方向にそれぞれ k, l 番目のものを、

代表点 (i, j, k, l) ( $-m \leq k \leq m, -n \leq l \leq n$ ) と呼ぶ。図5(b)からわかるように、代表点 (i, j, 0, 0) は格子点 (i, j) に等しい。

【0019】以下に、動きベクトル検出部15の具体的なブロック図を示した図6を用いて、動きベクトルを求める方法を説明する。

【0020】動きベクトル検出部15の入力は映像信号

であり、 $r$  フレームに 1 回 ( $r$  : 所定数) 入力されるよう に設定されているとする。ここで、ある時刻  $t$  の画像を第 0 フレーム目の画像とし、以後、時刻 ( $t + \tau$ ) の画像を第  $(30 + \tau)$  フレーム目の画像と呼ぶことにする。

【0021】今、第  $R$  フレーム目の画像が入力されたものとする。入力画像はまず、BPF23においてバンドパスフィルタに通される。ここで座標位置 ( $x, y$ ) での BPF 处理後の画像の値を  $I(x, y)$  とする。

【0022】一方、代表点値記憶部 24 は、 $r$  フレーム前すなわち第  $(R - r)$  フレーム目の BPF 处理後の画像の代表点の値が記憶されている。すなわち、代表点

$(i, j, k, l)$  の値

$$* \sum_k \sum_l \{ Y(i, j, k, l) - I(pos_x(i, 0) + g, pos_y(j, 0) + h) \}$$

【0025】が最小となる  $g, h$  を  $(2 \cdot G), (2 \cdot H)$  の範囲内  $(-G \leq g \leq G, -H \leq h \leq H)$  で探索することにより、動きベクトル  $(g, h)$  が求まる。

【0026】代表点値記憶部 24 の内容は、マッチング部 25 の処理が終了した後、更新される。具体的には、代表点位置記憶部 26 において記憶されている代表点の座標

$pos_x(i, j, k, l), pos_y(i, j, k, l)$

$1 \leq i \leq M, 1 \leq j \leq N, -m \leq k \leq m, -n \leq l \leq n$

を用いて、第  $R$  フレーム目の BPF 处理後の画像の代表点での値を記録する。

【0027】以上のようにして、入力された画像と  $r$  フレーム前の画像の 2 枚の画像から動きベクトルを求めることができる。

【0028】次に、カメラワークパラメタ推定部 16 において、動きベクトルからカメラワークパラメタを推定する方法を説明する。

【0029】動画像から推定できるカメラワークは、カメラの水平、垂直方向の変化（パンニング、チルティング）、カメラ画角の変化（ズーミング）、カメラの水平・垂直・前後の位置の変化（トラッキング、ブーミング、ドリーイング）などが考えられる。本実施例では簡単のため、パンニング、チルティング、ズーミングの 3 種類の操作を推定する方法を説明する。

【0030】まず、上記 3 種類のカメラワークによって、カメラの撮像面に投影された点がどのように移動するか考える。図 7 はカメラの撮像面と被写体の位置関係を示す図であり、カメラの空間の 3 次元座標を  $(x, y, z)$  で表し、撮像面上の 2 次元座標を  $(X, Y)$  で表している。また、カメラの位置を 3 次元座標の原点とし、カメラの光軸を  $z$  軸とする。撮像面は  $z = F$  ( $F$  : 焦点距離) に位置し、被写体の任意の点の座標  $u_1 = (x_1, y_1, z_1)$  が撮像面の  $U_1 = (X_1, Y_1)$  に投影されることを示している。ここで被写体の座標と撮像面上の座標との関係は、

\*  $Y(i, j, k, l) = I(pos_x(i, k), pos_y(j, l))$   
 $1 \leq i \leq M, 1 \leq j \leq N, -m \leq k \leq m, -n \leq l \leq n$   
 $pos_x(i, k)$  : 代表点  $(i, j, k, l)$  の  $x$  座標  
 $pos_y(i, k)$  : 代表点  $(i, j, k, l)$  の  $y$  座標  
 が記憶される。

【0023】マッチング部 25 は、BPF23 から BPF 处理後の画像  $I(x, y)$  を、代表点値記憶部 24 から  $r$  フレーム前の代表点の値  $Y(i, j, k, l)$  を入力し、代表点マッチングにより各格子点での動きベクトルを求め る。すなわち、格子点  $(i, j)$  に関して、

【0024】

【数 1】

【0031】

【数 2】

$$X_1 = F \cdot x_1 / z_1$$

$$Y_1 = F \cdot y_1 / z_1$$

【0032】で表せる。図 7 の座標を用いて、まずズーミングによる、被写体の座標の撮像面上の移動を考える。図 8(a)は、焦点距離の変化によって起こるズーミングを示したものである。同図のように焦点距離が  $F$  から  $F'$  に変化したとき、 $u_1$  の被写体の投影が  $U_1 = (X_1, Y_1)$  から  $U_2 = (X_2, Y_2)$  に移動する。

【0033】ただし、(数 2) から  $U_2$  は

$$U_2 = U_1 \cdot F' / F = f \cdot U_1$$

ただし、 $f = F' / F$  を満たす。

【0034】同様にして図 8(b)を用いてパンニング、チルティングの場合を考える。パンニング、チルティングはそれぞれカメラを  $y$  軸、 $x$  軸について回転する操作に等しい。同図のようにカメラが  $x$  軸について  $\theta_x$  だけ回転した場合、被写体の空間での座標  $u_1$  は  $u_3$  に移動する。ただし、 $u_3$  は(数 3)をみたす。

【0035】

【数 3】

$$u_3 = u_1 \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_x & -\sin \theta_x \\ 0 & \sin \theta_x & \cos \theta_x \end{bmatrix}$$

【0036】 $x$  に関する回転角  $\theta_x$  が十分小さいと仮定すると、移動後の撮像面上の座標  $U_3 = (X_3, Y_3)$  に対して(数 2)の関係から  
 $X_3 = X_1, Y_3 = Y_1 + F \cdot \theta_x$   
 の関係が導かれる。これを一般化すると、 $x$  軸、 $y$  軸に 対してともに回転するカメラ操作の場合、任意の座標の操作前後の関係は

$U_3 = U_1 + P$

ただし、 $P = (px, py)$

$px, py$ ：x軸、y軸の回転成分  
と表すことができる。

【0037】以上のことからズーミング、パンニング、チルティングを合成した一般的なカメラ操作に対して、カメラ操作前後の座標  $U_1 = (X_1, Y_1)$  、  $U' = (X', Y')$  は

$$U' = f \cdot U + P$$

を満たすことがわかる。以後  $f$  をズーム要素、 $P$  を回転ベクトルと呼ぶ。

【0038】従って、ズーム要素と回転ベクトルを求めることにより、カメラの操作量を推定することができる

ことがわかる。

【0039】以下に、動きベクトル検出部15で求めた

動きベクトルから、ズーム要素と回転ベクトルを推定す\*

$$f = \frac{\sum_{i,j} \langle U^{real} i, j, U_i, j \rangle - \langle \sum_i \sum_j U^{real} i, j, \sum_i \sum_j U_i, j \rangle / M / N}{\sum_{i,j} \langle U_i, j, U_i, j \rangle - \langle \sum_i \sum_j U_i, j, \sum_i \sum_j U_i, j \rangle / M / N}$$

$$P = (\sum_{i,j} U^{real} i, j - f \cdot \sum_{i,j} U_i, j) / M / N$$

【0042】と決まる。ただし、 $\langle \cdot, \cdot \rangle$  は内積を示す。従って、カメラワークパラメタ推定部16では、動きベクトル検出部15から動きベクトル  $v_{i,j}$  と格子点位置  $U_{i,j}$  を入力し、(数4)によって  $f, P$  を計算することにより、ズーミング、パンニング、チルティングの各カメラワークパラメタを推定することができる。

【0043】次に被写体情報検出部17の動作について説明する。被写体情報検出部17はビデオカメラで被写体をトラッキングしている状態において、被写体の位置や大きさ、色などの被写体情報を抽出するものである。すなわち、カメラワーク検出部14でパンニングを検出し、さらに動領域検出部18で動領域を検出できた場合※

$$|v_{i,j} - V_p| > \epsilon$$

【0046】によるカメラの動きベクトルとは異なる領域を抽出する。領域内特微量抽出部19では、動領域検出部18で検出した動領域から領域内の特微量として、重心位置と面積と色を抽出する。

【0047】次にフォーカス情報検出部20の動作について説明する。フォーカス情報検出部20は画像のピンボケ状態を検出するためのもので、画像の高周波成分の量をもとにしている。すなわち、画像がレンズの焦点ずれなどでぼけた状態では映像信号の高周波成分の値が小さくなる。このため、高域通過フィルタ21で画像の高周波成分を取り出し、平均値算出部22で画面全体、もしくは指定領域内の高周波成分の平均値を求める構成としている。

【0048】以上のようにして映像信号を処理することによって、代表画像を抽出するための情報を獲得することができる。本実施例では  $\gamma$  補正值、色温度、逆光や過

\*る方法を説明する。ここで、格子点  $(i, j)$  に関して、位置(2次元座標)を  $U_{i,j}$ 、動きベクトル検出部15で求められた動きベクトルを  $v_{i,j}$  とする。

【0040】今、ズーム要素  $f$ 、回転ベクトル  $P$  のカメラ操作が起こったとき、格子点  $(i, j)$  は  $U'_{i,j} (f, P) = f \cdot U_{i,j} + P$  の位置に移動するはずである。従って実際に起こったカメラ操作の  $f, P$  を推定するには、実際に移動した位置  $U^{real}_{i,j} = U_{i,j} + v_{i,j}$

との誤差

$E(f, P) = \sum (U'_{i,j} (f, P) - U^{real}_{i,j})^2$  が最小になるような  $f, P$  を求めればよい。誤差  $E$  は  $f, P$  に関して2次式なので、誤差  $E$  を最小とする  $f, P$  は一意に

【0041】

【数4】

【0042】

※に対して、領域内特微量抽出部19で動領域から領域内の特微量を抽出する。動領域検出部18における動作をさらに詳述する。

【0044】動領域検出部18には、動きベクトル検出部15で検出した画面内の  $M \cdot N$  個の格子点の動きベクトル  $v_{i,j}$  と、カメラワーク検出部14で検出したパンニングによるカメラの動きベクトル  $V_p$  が入力される。動領域検出部18では、(数5)を満たす格子点を抽出し、抽出した格子点の連結関係にもとづいて、パンニング

【0045】

【数5】

ただし、 $\epsilon$  は所定の値

順光状態、被写体光量などについては記載しなかったが、これらの情報も映像信号を処理することによって獲得することができる。代表画像を抽出するための情報を獲得した後、獲得した情報をもとに代表画像を抽出する構成と手法については静止画像抽出装置の第1の実施例と同様であり、説明は省略する。

【0049】以上説明した代表画像抽出情報を獲得する装置は、動画像記録媒体に代表画像を抽出するための情報が存在しない場合に、動画像記録媒体から読み込んだ映像信号をもとにして代表画像抽出情報を獲得するものであったが、同様にしてビデオカメラで撮影中に、撮像素子から取り込んだ映像信号をもとにして代表画像抽出情報を獲得することもできる。この場合の構成は図3と同様であり説明は省略するが、ビデオカメラに備えたセンサーによって検出可能な代表画像抽出情報が存在する場合は、映像信号から代表画像抽出情報を獲得する必要

がないことはいうまでもない。さらに、ビデオカメラで撮影中に獲得した代表画像抽出情報は、映像信号とともに動画像記録媒体に記録してもよい。

【0050】次に本発明の動画像記録装置の実施例の構成を図9に示す。本実施例はビデオカメラに静止画像抽出装置を備えた構成であり、ビデオカメラで撮影中に代表画像として抽出する静止画像を決定し、動画像記録媒体に映像信号を記録するとともに抽出する静止画像の情報を記録するものである。図9で本発明の動画像記録装置27は、カメラ操作情報獲得部28、画像処理情報獲得部29、撮影状態情報獲得部30、映像信号獲得部31、代表画像抽出情報評価部32、静止画像情報記録部33、映像信号記録部34からなる。以上の構成における各部の動作について以下で詳細に説明する。

【0051】カメラ操作情報獲得部28は、ビデオカメラで撮影した際の録画開始操作やズーム操作などの情報を獲得する部分である。録画開始操作情報は録画開始操作が行われた時点にフラグを立てて録画開始点を示す情報であり、ズーム操作情報はズーム倍率を表す情報である。録画開始操作情報もズーム操作情報とともに、ビデオカメラのボタン操作をもとに検出する。画像処理情報獲得部29は、撮像素子で撮像した映像信号を処理した情報を獲得する部分で、例えばフォーカス制御を行うために求めた映像信号の高周波成分の周波数や大きさ、あるいはフレーム間における輝度信号や色信号の差異を求めたフレーム間差分値、あるいは映像信号から求めた被写体領域の位置や大きさなどの情報、逆光や過順光の状態、さらにはγ補正値や色温度などを抽出する。撮影状態情報獲得部30は、カメラの撮影状態をセンサーで検出した情報を獲得する部分で、例えば角速度センサーによって検出したパンニングなどのカメラの動き情報、あるいは光量センサーによる被写体光量や絞り開度センサーで検出したレンズの絞り開度、あるいはレンズ位置検出によるフォーカス距離などを獲得する。

【0052】代表画像抽出情報評価部32は、カメラ操作情報に含まれる録画開始操作情報を検出し、次の録画開始操作情報を検出するまでの同一カット内の各画像に対して、カメラ操作情報の中の録画開始操作情報以外の情報と画像処理情報と撮影状態情報をもとに、各画像がカットの代表画像として適当な画像であるかどうかの評価を行う。代表画像として適当であると評価された画像に関して、静止画像情報記録部33を介して静止画像情報を記録媒体35の静止画像情報記録媒体に記録する。なお、代表画像抽出情報評価部32の動作については、後述の静止画像自動抽出方式の実施例で詳細に説明する。

【0053】以下では、静止画像情報記録部33で記録する静止画像情報についてさらに説明する。静止画像情報とは、代表画像抽出情報評価部32で代表画像として適当であると評価された静止画像そのもの、あるいは縮

小などの画像処理したもの、あるいは静止画像に対応する動画像記録媒体上の格納位置情報、あるいは静止画像に対応する動画像に付与したフラグなどである。

【0054】静止画像情報が静止画像そのもの、あるいは縮小した画像の場合は、映像信号獲得部31と映像信号記録部34によってカメラで撮影した動画像を記録する動画像記録媒体上の記録位置とは異なる位置、あるいは動画像を記録する動画像記録媒体とは異なる記録媒体に静止画像情報である画像を記録する。たとえば動画像記録媒体がビデオテープの場合は、静止画像情報の画像だけをまとめてテープの先頭部分、もしくはテープの終端部分に記録するか、あるいはテープとは別に備えたICメモリに静止画像情報だけを記録する。

【0055】静止画像情報が静止画像に対応する動画像記録媒体上の格納位置情報の場合は、映像信号獲得部31と映像信号記録部34によってカメラで撮影した動画像を記録する動画像記録媒体上の記録位置とは異なる位置、あるいは動画像を記録する動画像記録媒体とは異なる記録媒体に、静止画像情報である静止画像に対応する動画像記録媒体上の格納位置情報を記録する。

【0056】静止画像情報が静止画像に対応する動画像に付与したフラグの場合は、映像信号獲得部31と映像信号記録部34によってカメラで撮影した動画像を記録する動画像記録媒体上の記録位置と同じ位置に静止画像情報を記録する。すなわち、たとえば1フレーム単位で記録する映像信号の先頭部分に静止画像情報のフラグを記録する。

【0057】以上のようにしてビデオカメラで撮影した映像信号を動画像記録媒体に記録するとともに、撮影した動画像の中から代表画像を抽出し、代表画像の静止画像情報を記録媒体に記録する。これによって記録された静止画像情報を読みだして代表画像をディスプレイやプリンタに出力することが可能となる。

【0058】次に本発明における静止画像自動抽出方法の実施例を説明する。静止画像自動抽出方式は、図2の代表画像抽出情報評価部6および図9の代表画像抽出情報評価部32における処理方法である。

【0059】本発明の静止画像自動抽出方法は、録画開始操作をしてから録画終了操作をするまでの間に連続して撮影された動画像の中から、代表的な画像を静止画像として自動的に抽出するものである。ここで代表的な画像とは、撮影者の意図、撮影された画像の状態、被写体の状態をもとに評価して選びだした画像をいう。

【0060】撮影者の意図は、ズームやパンなどのカメラワークに反映される。すなわち、ズームインしているときは、注目している被写体が画面に存在する場合であり、重要な画像と考えられる。また、パンしているときは、ある場面から別の場面に移動している最中であり、重要ではないと考えられる。さらに、パンしている場合でも、トラッキングしている被写体が存在する場合は、

重要であると考えられる。このようにカメラワークから撮影者の意図を推定し、重要な部分を代表画像として抽出することが望ましい。

【0061】撮影された画像の状態とは、撮影時にフォーカス制御がうまく行われていない場合のぼけた画像状態や、アイリス制御が不適切な場合の過順光や逆光の状態、さらには $\gamma$ 補正が不適切な場合などの状態、あるいは絞りやフォーカスの調整中の状態などをいう。これらの画像状態は、ビデオカメラでの撮影時におけるフォーカス制御やアイリス制御の情報、あるいは $\gamma$ 補正值をもとに判断することができる。また、フォーカス制御やアイリス制御、 $\gamma$ 補正の情報がない場合であっても、映像信号を処理することによって求めることができる。これらの画像状態を評価して、代表画像としては画像状態が良好なものを抽出することが望ましい。

【0062】被写体の状態とは、撮影している被写体の位置や大きさなどの状態、撮影中にカメラの前を人が横切った場合などの障害物の存在状況、撮影中にフラッシュが光った場合などの状態、被写体にスポットライトが照射されている状態などをいう。被写体の位置や大きさに関しては、位置がカメラの中央で面積が大きいほうが望ましく、障害物は存在しない方が望ましい。また、フラッシュが光った場合の画像は代表画像として抽出しない方が望ましい。また、スポットライトが照射されている画像は注目画像であり、代表画像として抽出することが望ましい。ここで、被写体の位置や面積の検出方法に関しては第3の実施例において図3の被写体情報検出部17で説明している。また、フラッシュや障害物の検出方法に関しては、図3のフレーム間差分値検出部10で求めたフレーム間差分値をもとに検出可能である。すなわち、フラッシュはフレーム間差分値が急変するため、所定のしきい値をもとに検出できる。障害物の場合は、画面の中に障害物が入るときと画面から障害物が出るときにフレーム間差分値が変化するため、フレーム間差分値が所定のしきい値を越えてから、所定の時間内に再びフレーム間差分値がしきい値を越える場合は、障害物が画面内に存在するとして検出できる。また、スポットライトの照射は、被写体光量をもとに検出することができる。

【0063】以上のような代表画像を抽出するための知識に基づいて、以下では具体的な代表画像の抽出手法を説明する。本実施例の構成例を図10に示す。図10で36は重み付け加算部、37はゲート信号発生部、38はゲート部、39は最大値検出部である。重み付け加算部36にはズーム倍率と被写体情報が入力され、それぞれの信号に重みを付けて加算する。ここで、被写体情報はカメラで被写体をトラッキングしているときに得られる情報で、トラッキング時の被写体の位置と大きさをもとにしている。被写体情報は、被写体位置がカメラの中心に近いほど、また被写体の面積が大きいほど大きな値

になるようになる。ゲート部38は、ゲート信号発生部37のゲート信号をもとにスイッチのオン、オフを行う。最大値検出部39は、ゲート部38から入力される値の最大値を検出する。

【0064】ゲート信号発生部37は、パン信号と映像信号の高周波成分値とフレーム間差分値などをもとにゲート信号を発生する。ゲート信号の発生方法を図11に示す。図11(a)はパン信号で、パンニング中が0で、パンしていないときが1となる信号である。(b)

10 10は映像信号の高周波成分値であり、値が小さいほど画像がぼけた状態であることを意味する。(c)は(b)の信号をしきい値処理して2値化したものであり、しきい値以下の場合は0にしている。。(d)はフレーム間差分値である。(e)は(d)の信号をしきい値処理し、しきい値以上の場合は0にし、さらに(d)の信号がしきい値を越えてから所定時間内に再びしきい値を越えた場合は、0と0の間の区間も0にする処理を行う。すなわち、単独でフレーム間差分値が大きくなる場合は、フラッシュなどによる画像異常が発生したものと判断し、フレーム間差分値がしきい値以上である期間だけ0にする。しかし、前述したように障害物がカメラの前を通過する場合には、フレーム間差分値は複数のピークをとるため、障害物が画面の中に存在する期間はフレーム間差分値がしきい値以下であっても0にする。以上のようにして2値化した(a)(c)(e)の3信号の論理積をとることによって、ゲート信号を発生する。

【0065】以上のようにして図10で示した構成で評価値が最大値となる画像を求めるこによって、パンしている期間や画像がぼけた期間、さらにフラッシュや障害物が存在する期間を除いた中から、ズーム倍率が高く、被写体が画面中央に大きく写っている画像を代表画像として抽出することができる。なお、最大値検出部39で最大値を検出するのは、1カット全体であっても、または1カット内の複数の区間であってもよい。

【0066】なお、図10では重み付け加算部36とゲート信号発生部37とゲート部38の構成によって複数の入力からひとつの評価値を求めているが、この構成に限ったものではなく、ファジィ推論などのルールに基づいたものや、ニューラルネットワークによって求める構成も可能である。さらに、本実施例の構成には $\gamma$ 補正值や被写体光量、逆光や過順光状態、絞り開度、フォーカス距離に関する情報の処理を示さなかったが、これらの信号も同様にして利用することができる。すなわち、 $\gamma$ 補正值や絞り開度やフォーカス距離の値が変動しているとき、また逆光や過順光状態のときには代表画像として抽出しないようにゲート信号を発生させてもよい。また、被写体光量からスポットライトが照射されていることを検出して評価値を高くするようにしてもよい。

【0067】以上説明した静止画像自動抽出方式の実施例の構成は、撮影者が撮影開始操作をしてから撮影終了

操作をするまでの間に撮影された動画像のすべてに対し評価を行って代表画像を抽出しているが、撮影者が撮影開始操作をしてから所定時間経過後の画像から評価を行い、所定の条件を満足した時点の画像を代表画像として抽出する構成でもよい。以下ではこの構成の実施例について説明する。

【0068】本実施例の構成を図12に示す。図12で40はタイマー、41はゲート部、42は評価部である。タイマー40は、撮影開始操作が行われてからの経過時間を測定し、撮影開始から一定時間が経過した時点でゲート部41のゲートを開くようにゲート信号を発生する。評価部42はゲート部41を通過したフォーカスの高周波成分値と、フレーム間差分値が条件を満足しているかどうかの評価を行う。評価部42での条件は、フォーカスの高周波成分値が所定のしきい値以上で、かつフレーム間差分値が所定のしきい値以下であるという条件である。評価部42では条件が満足された時点でそれ以降の評価を中止し、条件が満足された時点の静止画像を代表画像として抽出する。なお、本実施例では評価に用いた信号はフォーカスの高周波成分値とフレーム間差分値のふたつだけであるが、パンやズームなどの他の信号を用いてもよい。

【0069】さらに静止画像自動抽出方式の別の実施例の構成を説明する。本実施例の構成を図13に示し、図12と同一のものには同一番号を付け説明は省略する。本実施例は、撮影開始操作から撮影終了操作までの間に撮影された画像の数に一定比率を乗じた画像数から評価を行い、所定の条件を満足した時点の画像を代表画像として抽出するものである。図13の構成では撮影開始操作から撮影終了操作までの間に撮影された画像に対し、中間フレーム以降の画像を評価するものである。このために中間フレーム検出部43において、撮影開始操作から撮影終了操作までに撮影されたフレームの中間フレームを検出し、中間フレームを検出した時点でゲート部41のゲートを開くようにゲート信号を発生する。ゲート部41および評価部42の動作は図12の構成と同様であり、説明は省略する。

#### 【0070】

【発明の効果】以上説明したように、本発明の動画像記録媒体は画像抽出情報を備えることによって、画像抽出情報をもとに動画像記録媒体に記録された動画像中の代表画像を容易に抽出することが可能となる。

【0071】また、本発明の静止画像抽出装置は画像抽出情報をもとにして、動画像の中から動画像の内容を十分に表現する静止画像を代表画像として抽出することが可能となり、動画像の内容を短時間で把握することができる。

【0072】さらに、本発明の動画像記録装置はカメラで撮影した映像信号を動画像記録媒体に記録するととも

に、撮影した動画像の中から代表画像を抽出し、代表画像の静止画像情報を記録媒体に記録することによって、記録された静止画像情報を読みだして代表画像をディスプレいやプリンタに高速に出力することが可能となる。

#### 【図面の簡単な説明】

【図1】本発明の実施例の動画像記録媒体を示す図

【図2】本発明の第1の実施例の静止画像抽出装置の構成を示す図

【図3】本発明の第2の実施例の静止画像抽出装置における代表画像情報抽出装置の構成を示す図

【図4】本発明の第2の実施例の静止画像自動抽出装置におけるフレーム間差分値検出部の他の構成を示すプロック図

【図5】本発明の第2の実施例の静止画像自動抽出装置における動きベクトル検出部における動きベクトルを検出する格子点を説明する図

【図6】本発明の第2の実施例の静止画像自動抽出装置における動きベクトル検出部の構成を示す図

【図7】カメラの撮像面と被写体の位置関係を示す図

【図8】ズーミングとチルティングでのカメラの撮像面と被写体の位置関係を示す図

【図9】本発明の実施例の動画像記録装置の構成を示す図

【図10】本発明の第1の実施例の静止画像自動抽出方法の代表画像抽出情報評価部の構成を示す図

【図11】本発明の第1の実施例の静止画像自動抽出方法のゲート信号発生部の動作を説明する図

【図12】本発明の第2の実施例の静止画像自動抽出方法の代表画像抽出情報評価部の構成を示す図

【図13】本発明の第3の実施例の静止画像自動抽出方法の代表画像抽出情報評価部の構成を示す図

#### 【符号の説明】

1 再生信号入力部

2、28 カメラ操作情報獲得部

3、29 画像処理情報獲得部

4、30 撮影状態情報獲得部

5、31 映像信号獲得部

6、32 代表画像抽出情報評価部

7 代表画像管理部

8 代表画像記憶部

9 出力装置

33 静止画像情報記録部

36 重み付け加算部

37 ゲート信号発生部

38、41 ゲート部

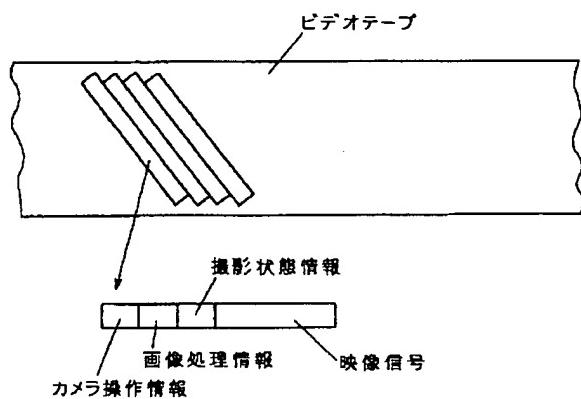
39 最大値検出部

40 タイマー

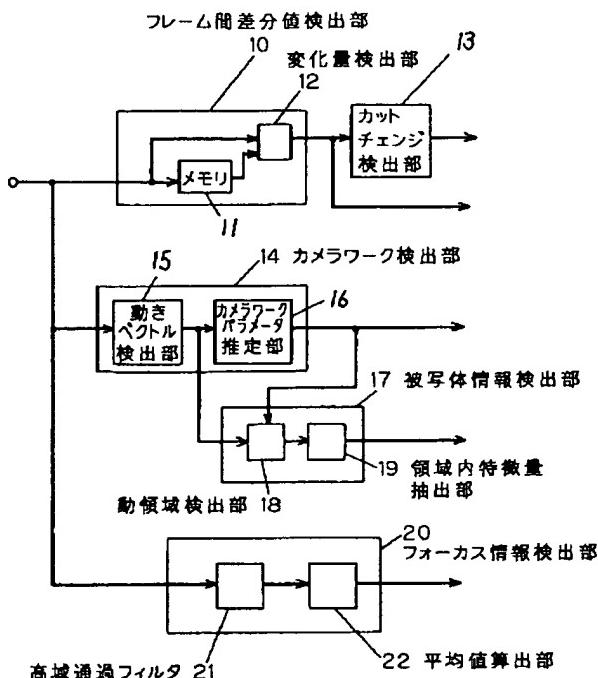
42 評価部

43 中間フレーム検出部

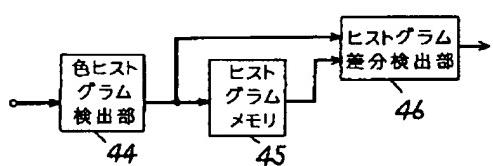
【図1】



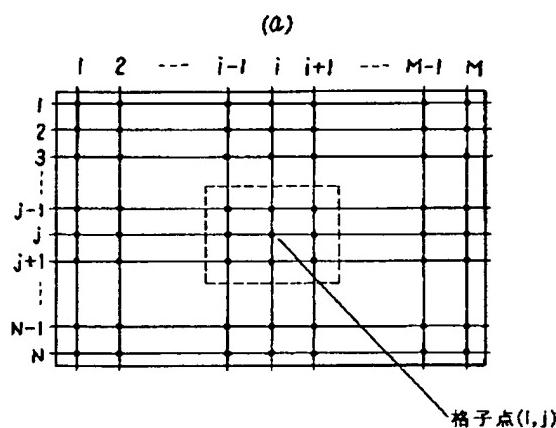
【図3】



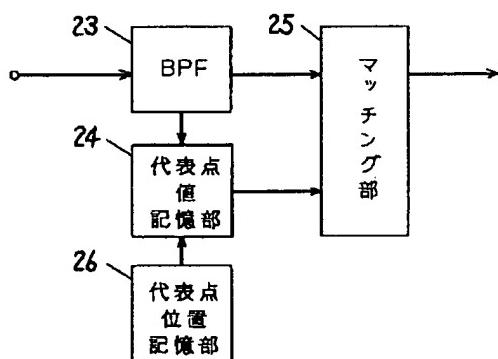
【図4】



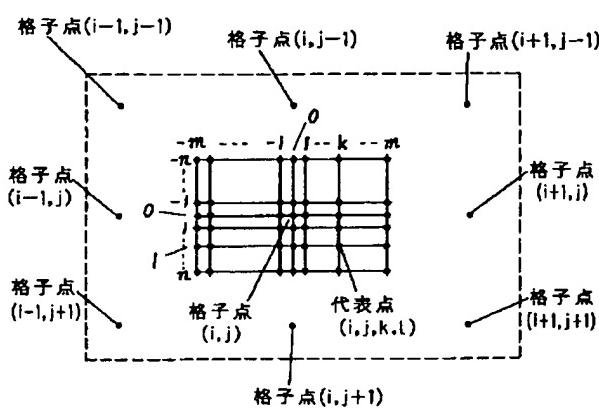
【図5】



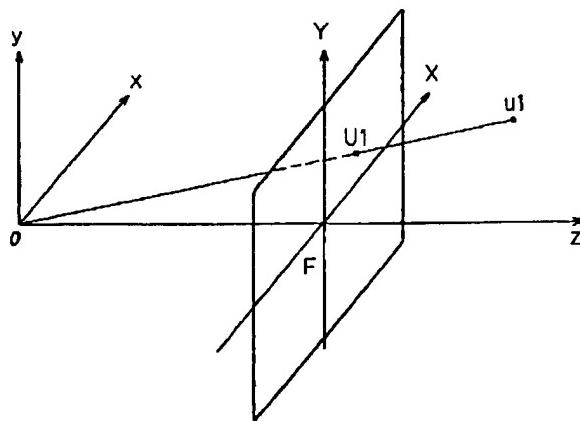
【図6】



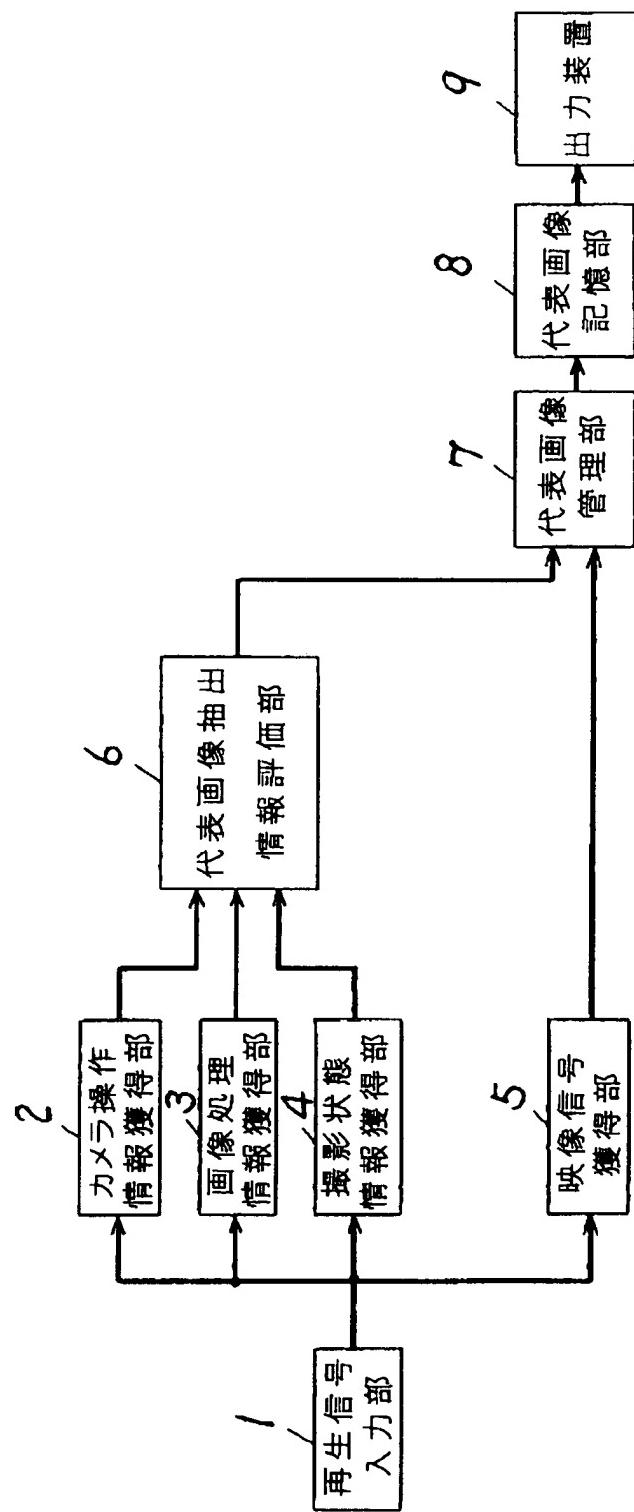
(b)



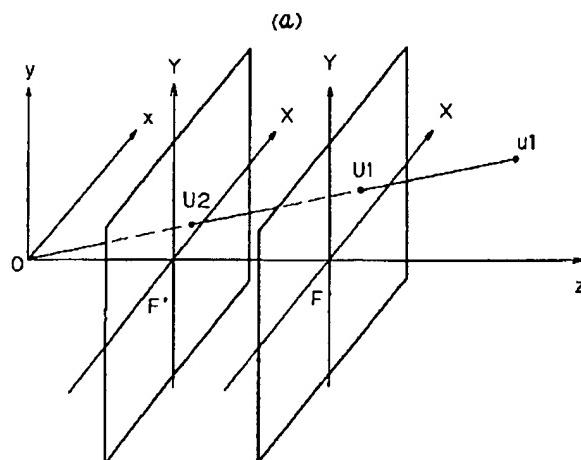
【図7】



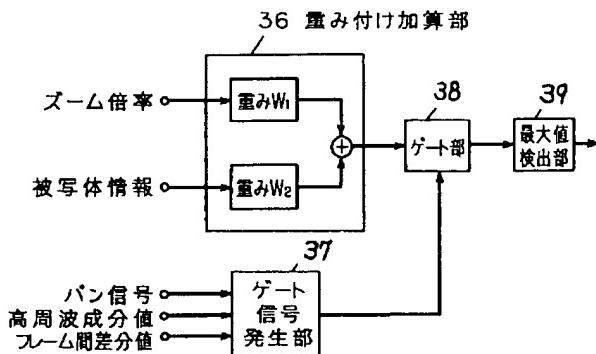
【図2】



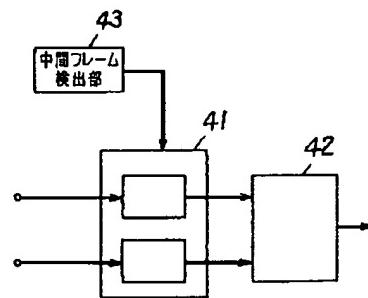
【図8】



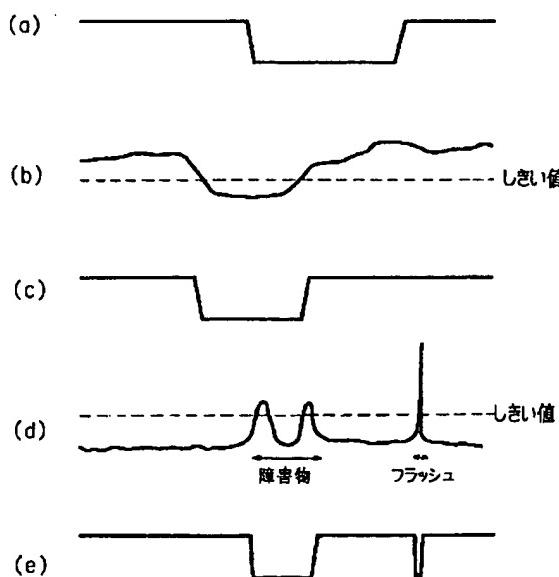
【図10】



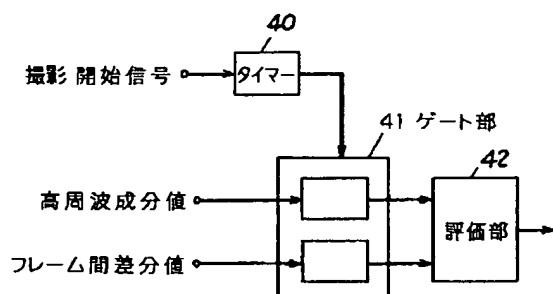
【図13】



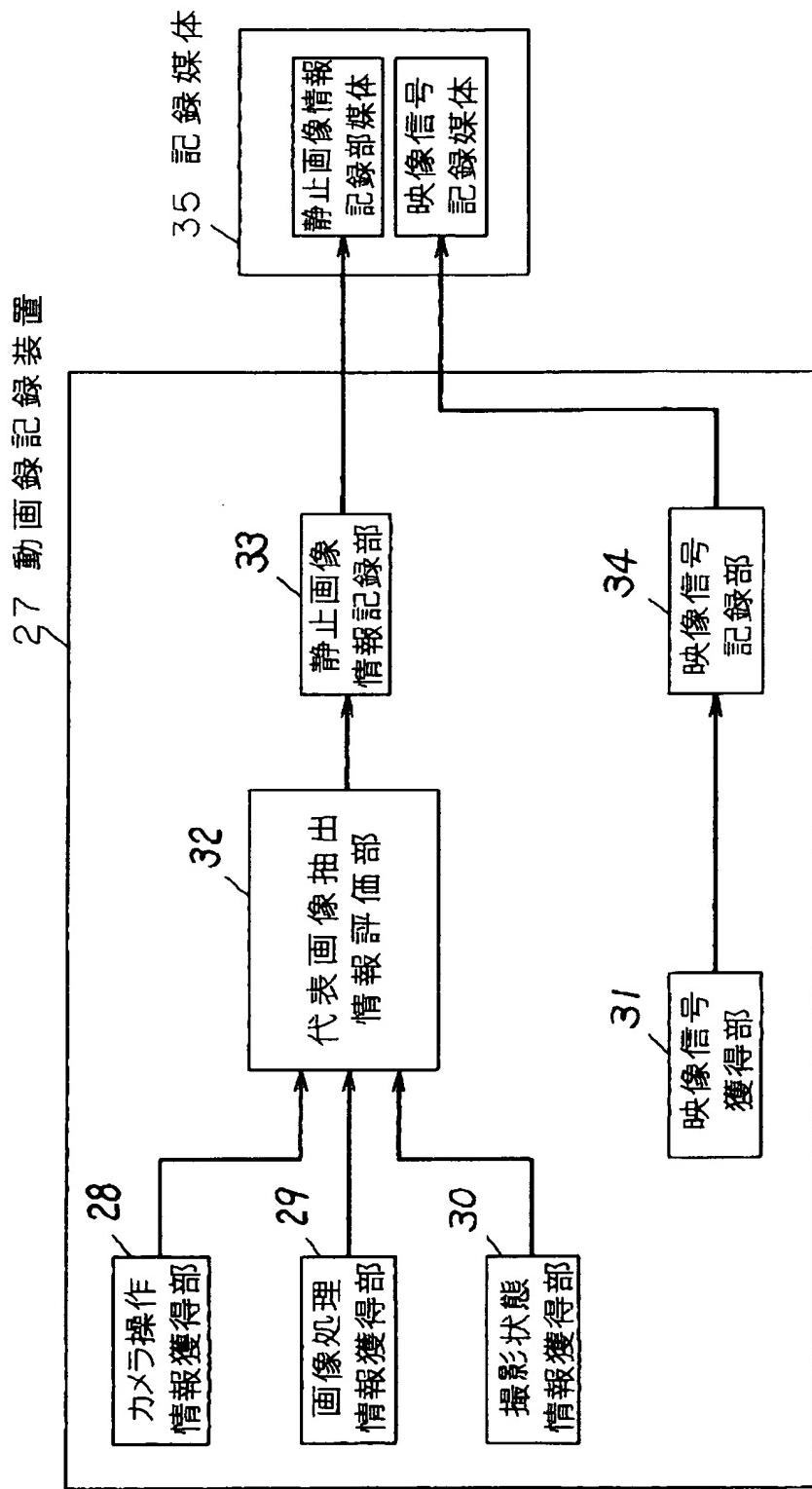
【図11】



【図12】



【図9】



フロントページの続き

(51) Int.Cl.<sup>6</sup> 識別記号 庁内整理番号 F I 技術表示箇所  
H 04 N 5/232 Z  
5/91